Linear Algebra and its Applications 430 (2009) 1292–1312

Contents lists available at ScienceDirect

Linear Algebra and its Applications

journal homepage: www.elsevier.com/locate/laa

Normal form for linear systems with respect to its vector relative degree

Markus Mueller

Institute of Mathematics, Technical University Ilmenau, Weimarer Straße 25, 98693 Ilmenau, DE, Germany

ARTICLE INFO

Article history:
Received 28 January 2008
Accepted 20 October 2008
Available online 10 December 2008

Submitted by V. Mehrmann

AMS classification:
93B10
34C20

Keywords:
Linear systems
MIMO-systems
Vector relative degree
Normal form

ABSTRACT

For multi-input multi-output (MIMO) linear systems with existing vector relative degree a normal form is constructed. This normal form is not only structural simple but allows to characterize the system's zero dynamics for the design of feedback controllers. A characterization of the zero dynamics in terms of the normal form is given.

1. Introduction

In the present work, linear systems with m inputs and m outputs of the form

$$\left.\begin{aligned} \dot{x} &= Ax + \underbrace{[b_1^{(n)}, \ldots, b_m^{(n)}]}_{=B} \underbrace{\begin{pmatrix} u_1 \\ \vdots \\ u_m \end{pmatrix}}_{=u} \\ \underbrace{\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}}_{=y} &= \underbrace{\begin{bmatrix} c_{(n)}^1 \\ \vdots \\ c_{(n)}^m \end{bmatrix}}_{=C} x \end{aligned}\right\} \tag{1.1}$$

E-mail address: markus.mueller@tu-ilmenau.de

0024-3795/$ - see front matter © 2008 Elsevier Inc. All rights reserved.
doi:10.1016/j.laa.2008.10.014

Nomenclature

$\left[l_1^{(n)},\ldots,l_m^{(n)}\right] = L \in \mathbb{R}^{n\times m}$	where $l_i^{(n)} \in \mathbb{R}^n$ denotes the ith column of L and the superscript (n) remarks the dimension of the vector
$[(l_{(m)}^1)^T,\ldots,(l_{(m)}^n)^T]^T = L \in \mathbb{R}^{n\times m}$	where $l_{(m)}^j \in \mathbb{R}^{1\times m}$ denotes the jth row of L and the subscript (m) remarks the dimension of the row-vector
$e_k^{(n)} := [0_{1\times(k-1)}, 1, 0_{1\times(n-k)}]^T$	the kth row unit vector in $\mathbb{R}^n$
$e_{(m)}^k := [0_{1\times(k-1)}, 1, 0_{1\times(m-k)}]$	the kth row unit vector in $\mathbb{R}^{1\times m}$
$0_{n\times m} \in \mathbb{R}^{n\times m}$	the 0-matrix of dimension $n\times m$
$\mathcal{X}_{n\times m} \in \mathbb{R}^{n\times m}$	an arbitrary matrix of dimension $n\times m$; note that the use of this symbol implicates that the specific entries of the matrix are not important but only the dimension
$I_n \in \mathbb{R}^{n\times n}$	the identity matrix of dimension $n\times n$
$\mathcal{C}^m([0,\infty)\to\mathbb{R}^n)$	the set of m-times continuously differentiable maps from $[0,\infty)$ to $\mathbb{R}^n$
$\mathcal{C}_{\mathrm{pw}}([0,\infty)\to\mathbb{R}^m)$	the set of piecewise continuous maps from $[0,\infty)$ to $\mathbb{R}^m$

are considered, where $n,m \in \mathbb{N}$ with $m \leqslant n$ and $A \in \mathbb{R}^{n\times n}, B, C^T \in \mathbb{R}^{n\times m}$.

It is well known that a linear single-input single-output (SISO) system (1.1), i.e. $m=1$, has relative degree $r \in \mathbb{N}$ if, and only if, r is the least number of times one has to differentiate the output to have the input appear explicitly. In other words (see [11, Definition 2, 5, Definition 2.2]): defining, for a SISO-system (1.1) $H_0(x) := Cx$ and $H_{k+1}: \mathbb{R}^n \times \mathbb{R}^{k+1} \to \mathbb{R}, (x,u_0,\ldots,u_k) \mapsto \frac{\partial H_k}{\partial x}(Ax+Bu_0) + \sum_{j=0}^{k-1}\frac{\partial H_k}{\partial u_j}u_{j+1}$, for $k \in \mathbb{N}$, then $y^{(k)}(t) = H_k(x(t),u(t),\ldots,u^{(k-1)}(t))$, and (1.1) has relative degree $r\in\mathbb{N}$ if, and only if:

(i) $\forall\, k = 1,\ldots,r-1\ \forall\, i = 0,\ldots,k-1\ \forall\,(x,u_0,\ldots,u_{k-1}) \in \mathbb{R}^n\times\mathbb{R}^k\ :\ \frac{\partial H_k}{\partial u_i}(x,u_0,\ldots,u_{k-1}) = 0,$

(ii) $\forall\,(x,u_0,\ldots,u_{r-1}) \in \mathbb{R}^n\times\mathbb{R}^r\ :\ \frac{\partial H_r}{\partial u_0}(x,u_0,\ldots,u_{r-1}) \neq 0.$

In case of MIMO-system (1.1), for every $(i,j) \in \{1,\ldots,m\}\times\{1,\ldots,m\}$, one can consider the SISO-system relating input u_j to output y_i given by

$$\left.\begin{aligned} \dot{x} &= Ax + b_j^{(n)}u_j \\ y_i &= c_{(n)}^i x, \qquad i,j \in \{1,\ldots,m\} \end{aligned}\right\}. \tag{1.2}$$

Let $r_{i,j} \in \mathbb{N}$ be the relative degree of (1.2). Then, for $i \in \{1,\ldots,m\}$, $r_i := \min_{j\in\{1,\ldots,m\}} r_{i,j}$ is the least number one has to differentiate the ith output to have at least one of the m inputs appear explicitly in the sense as above. The vector $(r_1,\ldots,r_m) \in \mathbb{N}^{1\times m}$ is called the vector relative degree of the MIMO-system (1.1) if, for all $j \in \{1,\ldots,m\}$, the rows $c_{(n)}^j A^{r_j-1}B$ are linearly independent, see Definition 2.1(a).

Isidori [9] presents a local definition of the vector relative degree for nonlinear MIMO-systems.

Liberzon et al. [11] give a generalization of the relative degree for time-invariant nonlinear systems which is extended in [5] for time-varying linear and nonlinear systems. However in these papers only SISO-systems and MIMO-systems with strict relative degree (see Definition 2.1(c)) are considered.

The relative degree of a system leads to a normal form. For linear SISO-systems one can construct an invertible matrix $U \in \mathbb{R}^{n\times n}$ such that the coordinate transformation $\binom{\xi}{\eta} = Ux$ converts a linear system

$$\dot{x} = Ax + bu, \quad y = cx \tag{1.3}$$

with $A \in \mathbb{R}^{n\times n}$ and $b, c^T \in \mathbb{R}^n$, which has relative degree $r \in \mathbb{N}$, into

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\xi\\ \eta\end{pmatrix} = \left[\begin{array}{ccccc|c} 0 & 1 & & 0 & & 0 \\ \vdots & \ddots & \ddots & & & \vdots \\ 0 & \cdots & 0 & 1 & & 0 \\ R_1^1 & \cdots & & R_r^1 & & S^1 \\ \hline P_1 & 0 & \cdots & 0 & & Q \end{array}\right]\begin{pmatrix}\xi\\ \eta\end{pmatrix} + \left[\begin{array}{c} 0 \\ \vdots \\ 0 \\ cA^{r-1}b \\ \hline 0 \end{array}\right] u \tag{1.4}$$
$$y = [1, 0, \ldots, 0]\begin{pmatrix}\xi\\ \eta\end{pmatrix},$$

where $R_1^1, \ldots, R_r^1 \in \mathbb{R}, S^1 \in \mathbb{R}^{1\times n-r}, P_1 \in \mathbb{R}^{n-r}$ and $Q \in \mathbb{R}^{(n-r)\times(n-r)}$ can be presented explicitly in terms of the system matrices A, b and c, see [7]. This result is implicitly contained in [9, Chapter 4.1].

The Byrnes–Isidori normal form for nonlinear and linear SISO-systems, introduced in [1], is widely used in control theory for the design of local and global feedback stabilization of nonlinear systems [2–4], for the design of adaptive observers [12], for the design of adaptive controllers for linear systems [8,6], to name but a few applications. Thus a construction of a normal form for linear MIMO-systems will assist the design of controllers and observers for linear MIMO-systems.

Isidori [9, Chapter 5] presents a local normal form for nonlinear MIMO-systems. In [10, Chapter 11] a proof is given to specify the diffeomorphism to produce the normal form in terms of the system data of the nonlinear system. Moreover, for nonlinear systems that satisfy certain assumptions, namely commutativity of certain vector fields which is automatically satisfied for linear systems, Isidori [10, Proposition 11.5.2] gives a normal form which coincides with the normal form for linear systems given in the present work. However, the corresponding results for linear systems cannot be found in literature. One could translate the nonlinear results for linear systems, but the machinery for nonlinear systems, e.g. Lie-derivatives of smooth functions, is not necessary to prove the linear results. The proof given in the present work only uses standard linear algebra and is therefore independent of the nonlinear results. The transformation matrix is given in terms of the linear system matrices A, B and C and leads to "many zeros and ones" in the normal form and allows to read off the zero dynamics very easily; the reader will find that the normal form (2.1) for linear MIMO-systems is, roughly speaking, structured as a "diagonal form of m copies of SISO normal forms (1.4)". Furthermore, the matrices of the normal form and transformation will be characterized explicitly by the system matrices.

The present paper is structured as follows. In Section 2, the main results, i.e. the normal form for linear MIMO-systems is presented and the system's zero dynamics is characterized. Furthermore, the inverse of the system [9, Chapter 5.1] is presented. Section 3 contains all the proofs.

2. Normal form and zero dynamics

Consider, for $n, m \in \mathbb{N}$ with $m \leqslant n$ and $A \in \mathbb{R}^{n\times n}, B, C^T \in \mathbb{R}^{n\times m}$, a linear system (1.1), that is a linear system with m-dimensional input u and output y of form

$$\dot{x} = Ax + Bu,$$
$$y = Cx.$$

For linear MIMO-systems the vector relative degree is defined as follows.

Definition 2.1. A linear system (A,B,C) of form (1.1) has

(a) *(vector) relative degree* $r=(r_1,\ldots,r_m)\in\mathbb{N}^{1\times m}$ if, and only if:

(i) $\forall\, j\in\{1,\ldots,m\}\ \forall\, k\in\{0,\ldots,r_j-2\}\ :\ c^j_{(n)}A^kB=0_{1\times m}$,

(ii) $\operatorname{rk}\begin{bmatrix} c^1_{(n)}A^{r_1-1}B\\ c^2_{(n)}A^{r_2-1}B\\ \vdots\\ c^m_{(n)}A^{r_m-1}B\end{bmatrix}=m.$

(b) *Ordered (vector) relative degree* $r=(r_1,\ldots,r_m)\in\mathbb{N}^{1\times m}$ if, and only if, (1.1) has (vector) relative degree $r=(r_1,\ldots,r_m)$ with $r_1\geqslant r_2\geqslant\cdots\geqslant r_m$.

(c) *Strict relative degree* $\varrho\in\mathbb{N}$ if, and only if, (1.1) has (vector) relative degree $r=(r_1,\ldots,r_m)\in\mathbb{N}^{1\times m}$ with $\varrho=r_1=r_2=\cdots=r_m$.

Remark 2.2. (i) Note that Definition 2.1(a) coincides with the definition of the vector relative degree for nonlinear MIMO-systems, see [9, Chapter 5.1].

(ii) The linear independence of the rows $c^j_{(n)}A^{r_j-1}B$, although a quite restrictive requirement, is significant for the construction of a coordinate transformation and with it the normal form. For systems that do not satisfy both conditions in Definition 2.1(a) the vector relative degree does not exist and thus one cannot construct the normal form (2.1)–(2.2).

(iii) Note that in literature sometimes the relative degree is called uniform instead of strict.

The following lemma shows that the assumption of ordered vector relative degree is not restrictive.

Lemma 2.3. *Let (A,B,C) be a linear system of form (1.1) with vector relative degree $r=(r_1,\ldots,r_m)\in\mathbb{N}^{1\times m}$. Then there exists a permutation matrix $P\in\mathbb{R}^{m\times m}$ such that the system (A,B,PC) has ordered vector relative degree $rP=(\tilde r_1,\ldots,\tilde r_m)$.*

The following theorem presents a normal form for linear systems (A,B,C) of form (1.1) with ordered vector relative degree. The normal form has similar structural properties as the normal form for linear SISO-systems and linear MIMO-systems with strict relative degree, respectively, see (1.4).

Theorem 2.4. (i) *Consider a linear system (A,B,C) of form (1.1) with ordered vector relative degree $r=(r_1,\ldots,r_m)\in\mathbb{N}^{1\times m}$. Set $r^s:=\sum_{j=1}^m r_j$. Then there exists an invertible matrix $U\in\mathbb{R}^{n\times n}$ such that the coordinate transformation*

$$\begin{pmatrix}\xi\\ \eta\end{pmatrix}:=Ux,\quad \xi(t)=(y_1(t),\ldots,y_1^{(r_1-1)}(t)\,|\cdots|\,y_m(t),\ldots,y_m^{(r_m-1)}(t))^T\in\mathbb{R}^{r^s},\quad \eta(t)\in\mathbb{R}^{n-r^s}$$

converts (A,B,C) into

$$\left.\begin{aligned}\tfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\xi\\ \eta\end{pmatrix}&=\widetilde A\begin{pmatrix}\xi\\ \eta\end{pmatrix}+\widetilde Bu\\ y&=\widetilde C\begin{pmatrix}\xi\\ \eta\end{pmatrix}\end{aligned}\right\},\tag{2.1}$$

where

$$\left[\begin{array}{c|c}\widetilde{A} & \widetilde{B}\\ \hline \widetilde{C} & 0\end{array}\right] =$$

$$\left[\begin{array}{c|c}
\left[\begin{array}{ccc|ccc|c|ccc|c}
0 & 1 & 0 & 0 & \dots & 0 & & 0 & \dots & 0 & 0_{1\times(n-r^s)}\\
\vdots & \ddots & \ddots & \vdots & & \vdots & \dots & \vdots & & \vdots & \vdots\\
0 & \dots\ 0 & 1 & 0 & \dots & 0 & & 0 & \dots & 0 & 0_{1\times(n-r^s)}\\
R^1_{1,1} & \dots & R^1_{1,r_1} & R^1_{2,1} & \dots & R^1_{2,r_2} & & R^1_{m,1} & \dots & R^1_{m,r_m} & S^1\\ \hline
0 & \dots & 0 & 0 & 1 & 0 & & 0 & \dots & 0 & 0_{1\times(n-r^s)}\\
\vdots & & \vdots & \vdots & \ddots & \ddots & & \vdots & & \vdots & \vdots\\
0 & \dots & 0 & 0 & \dots\ 0 & 1 & & 0 & \dots & 0 & 0_{1\times(n-r^s)}\\
R^2_{1,1} & \dots & R^2_{1,r_1} & R^2_{2,1} & \dots & R^2_{2,r_2} & & R^2_{m,1} & \dots & R^2_{m,r_m} & S^2\\ \hline
 & \vdots & & & & & \ddots & & & & \vdots\\ \hline
0 & \dots & 0 & 0 & \dots & 0 & & 0 & 1 & 0 & 0_{1\times(n-r^s)}\\
\vdots & & \vdots & \vdots & & \vdots & & \vdots & \ddots & \ddots & \vdots\\
0 & \dots & 0 & 0 & \dots & 0 & & 0 & \dots\ 0 & 1 & 0_{1\times(n-r^s)}\\
R^m_{1,1} & \dots & R^m_{1,r_1} & R^m_{2,1} & \dots & R^m_{2,r_2} & & R^m_{m,1} & \dots & R^m_{m,r_m} & S^m\\ \hline
P_1 & 0\ \dots & 0 & P_2 & 0\ \dots & 0 & \dots & P_m & 0\ \dots & 0 & Q
\end{array}\right]
&
\left[\begin{array}{c}
0_{1\times m}\\ \vdots\\ 0_{1\times m}\\ c^1_{(n)}A^{r_1-1}B\\ \hline
0_{1\times m}\\ \vdots\\ 0_{1\times m}\\ c^2_{(n)}A^{r_2-1}B\\ \hline
\vdots\\ \hline
0_{1\times m}\\ \vdots\\ 0_{1\times m}\\ c^m_{(n)}A^{r_m-1}B\\ \hline
0_{(n-r^s)\times m}
\end{array}\right]\\ \hline
\left[\begin{array}{cccc|cccc|c|cccc|c}
1 & 0 & \dots & 0 & 0 & \dots & & 0 & & 0 & \dots & & 0 & \\
0 & \dots & & 0 & 1 & 0 & \dots & 0 & & \vdots & & & \vdots & \\
\vdots & & & \vdots & 0 & \dots & & 0 & \dots & & & & & 0_{m\times(n-r^s)}\\
 & & & & \vdots & & & \vdots & & 0 & \dots & & 0 & \\
0 & \dots & & 0 & 0 & \dots & & 0 & & 1 & 0 & \dots & 0 &
\end{array}\right] & 0
\end{array}\right] \tag{2.2}$$

and $R^j_{i,k} \in \mathbb{R}$, *for* $i,j \in \{1,\dots,m\}$ *and* $k \in \{1,\dots,r_i\}, S^1,\dots,S^m \in \mathbb{R}^{1\times(n-r^s)}, P_1,\dots,P_m \in \mathbb{R}^{n-r^s}$ *and* $Q \in \mathbb{R}^{(n-r^s)\times(n-r^s)}$.

(ii) *Following, the entries of the matrices in* (2.2) *and the entries of the transformation matrix U are expressed explicitly in terms of the system matrices A, B and C : Set*

$$m_i := \#\{r_j \mid r_j \geqslant i, \quad j \in \{1,\dots,m\}\}, \ i \in \{1,\dots,r_1\}, \tag{2.3}$$

the number of r_j*'s,* $j \in \{1,\dots,m\}$*, such that* $r_j \geqslant i$*, and define*

$$\Gamma := \begin{bmatrix} c^1_{(n)}A^{r_1-1}B\\ \vdots\\ c^m_{(n)}A^{r_m-1}B \end{bmatrix} \in \mathbb{R}^{m\times m}, \tag{2.4}$$

$$\mathcal{C} := \begin{bmatrix} c_{(n)}^{1} \\ \vdots \\ c_{(n)}^{1} A^{r_1 - 1} \\ \hline c_{(n)}^{2} \\ \vdots \\ c_{(n)}^{2} A^{r_2 - 1} \\ \hline \vdots \\ \hline c_{(n)}^{m} \\ \vdots \\ c_{(n)}^{m} A^{r_m - 1} \end{bmatrix} \in \mathbb{R}^{r^s \times n}, \tag{2.5}$$

$$\mathcal{B} := \left[B\Gamma^{-1}[e_1^{(m)}, \ldots, e_{m_1}^{(m)}], \quad AB\Gamma^{-1}[e_1^{(m)}, \ldots, e_{m_2}^{(m)}], \ldots, A^{r_1 - 1}B\Gamma^{-1}[e_1^{(m)}, \ldots, e_{m_{r_1}}^{(m)}] \right] \in \mathbb{R}^{n \times r^s}, \tag{2.6}$$

$$\mathcal{V} \in \mathbb{R}^{n \times (n - r^s)} \ : \ \operatorname{im} \mathcal{V} = \ker \mathcal{C}, \quad \text{and} \quad \operatorname{rk} \mathcal{V}^T \mathcal{V} = n - r^s, \tag{2.7}$$

$$\widehat{U} := \begin{bmatrix} \mathcal{C} \\ \mathcal{N} \end{bmatrix} \in \mathbb{R}^{n \times n} \quad \text{and} \quad \mathcal{N} := (\mathcal{V}^T \mathcal{V})^{-1} \mathcal{V}^T [I - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}], \tag{2.8}$$

$$T_i := \left[0_{(r_i + n - r^s) \times \left(\sum_{j=1}^{i-1} r_j\right)} \ \middle| \ \begin{matrix} I_{r_i} \\ 0_{(n - r^s) \times r_i} \end{matrix} \ \middle| \ 0_{(r_i + n - r^s) \times \left(\sum_{j=i+1}^{m} r_j\right)} \ \middle| \ \begin{matrix} 0_{r_i \times (n - r^s)} \\ I_{n - r^s} \end{matrix} \right]$$
$$\in \mathbb{R}^{(r_i + n - r^s) \times n}, \tag{2.9}$$

$$\widehat{\mathcal{C}}_i := [I_{r_i}, 0_{r_i \times (n - r^s)}] \in \mathbb{R}^{r_i \times (r_i + n - r^s)}, \tag{2.10}$$

$$\widehat{\mathcal{B}}_i := \left[e_{r_i}^{(r_i + n - r^s)}, (T_i \widehat{U} A \widehat{U}^{-1} T_i^T) e_{r_i}^{(r_i + n - r^s)}, \ldots, (T_i \widehat{U} A \widehat{U}^{-1} T_i^T)^{r_i - 1} e_{r_i}^{(r_i + n - r^s)} \right]$$
$$\in \mathbb{R}^{(r_i + n - r^s) \times r_i}, \tag{2.11}$$

$$\widehat{\mathcal{N}}_i := \left[0_{(n - r^s) \times r_i}, I_{n - r^s} \right] \left[I_{r_i + n - r^s} - \widehat{\mathcal{B}}_i (\widehat{\mathcal{C}}_i \widehat{\mathcal{B}}_i)^{-1} \widehat{\mathcal{C}}_i \right] \in \mathbb{R}^{(n - r^s) \times (r_i + n - r^s)}, \tag{2.12}$$

$$\widehat{U}_i := \left[\begin{array}{c|c} I_{r^s} & 0_{r^s \times (n - r^s)} \\ \hline 0_{(n - r^s) \times \left(\sum_{j=1}^{i-1} r_j\right)}, \ \widehat{\mathcal{N}}_i \begin{bmatrix} I_{r_i} \\ 0_{(n - r^s) \times r_i} \end{bmatrix}, \ 0_{(n - r^s) \times \left(\sum_{j=i+1}^{m} r_j\right)} & I_{n - r^s} \end{array} \right] \in \mathbb{R}^{n \times n} \tag{2.13}$$

for $i \in \{1, \ldots, m\}$, and finally

$$U := \widehat{U}_m \cdot \widehat{U}_{m-1} \cdots \widehat{U}_1 \cdot \widehat{U}. \tag{2.14}$$

Then, for $i, j \in \{1, \ldots m\}$, the entries in (2.1) *are given by*

$$[R_{j,1}^i, \ldots, R_{j,r_j}^i] = \left[c_{(n)}^i A^{r_i} \mathcal{B}(\mathcal{C}\mathcal{B})^{-1} \begin{bmatrix} 0_{\left(\sum_{\mu=1}^{j-1} r_\mu\right) \times r_j} \\ I_{r_j} \\ 0_{\left(\sum_{\mu=j+1}^{m} r_\mu\right) \times r_j} \end{bmatrix} \right.$$
$$\left. + \, c_{(n)}^i A^{r_i} \mathcal{V} [0_{(n - r^s) \times r_j}, I_{n - r^s}] \widehat{\mathcal{B}}_j (\widehat{\mathcal{C}}_j \widehat{\mathcal{B}}_j)^{-1} \right], \tag{2.15}$$

$$S^i = c_{(n)}^i A^{r_i} \mathcal{V}, \tag{2.16}$$

$$[P_i, 0, \ldots, 0] = \widehat{\mathcal{N}}_i (T_i \widehat{U} A \widehat{U}^{-1} T_i^T) \widehat{\mathcal{B}}_i (\widehat{\mathcal{C}}_i \widehat{\mathcal{B}}_i)^{-1}, \tag{2.17}$$

$$Q = \mathcal{N} A \mathcal{V} \overset{(2.8)}{=} (\mathcal{V}^T \mathcal{V})^{-1} \mathcal{V}^T [I - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}] A \mathcal{V}. \tag{2.18}$$

Next the definition of the zero dynamics of a linear system (A, B, C) of form (1.1) is given, compare, for example, [5, Definition 4.1]. The formal definition is very similar to Isidori [9, pp. 163–164] but is not based on the normal form. Furthermore, exponential stability of linear systems and exponential stability of the zero dynamics of a linear system is defined.

Definition 2.5. Introduce the *behaviour* of a linear system (A, B, C) of form (1.1):

$$\mathfrak{B}(A,B,C) := \Big\{(x,u,y) \in \mathcal{C}^1([0,\infty) \to \mathbb{R}^n) \times \mathcal{C}_{\mathrm{pw}}([0,\infty) \to \mathbb{R}^m) \quad \times \ \mathcal{C}^1([0,\infty) \to \mathbb{R}^m) \Big| \ (x,u,y) \text{ solves } (1.1)\Big\}.$$

(i) The *zero dynamics* of a linear system (A, B, C) of form (1.1) are defined as the real vector space of trajectories

$$\mathcal{ZD}(A,B,C) := \{(x,u,y) \in \mathfrak{B}(A,B,C) \mid y \equiv 0 \text{ on } [0,\infty)\}.$$

(ii) A linear system $\dot{x} = Ax$, for $A \in \mathbb{R}^{n\times n}$, is called *exponentially stable* on $[0,\infty)$ if, and only if

$$\exists M, \lambda > 0 \ \forall t \geqslant 0 \ : \ \|x(t)\| \leqslant Me^{-\lambda t}\|x(0)\|$$

for all solutions x of $\dot{x} = Ax$. A is called Hurwitz if, and only if, $\dot{x} = Ax$ is exponentially stable.

(iii) The *zero dynamics* of a linear system (A, B, C) of form (1.1) are called *exponentially stable* if, and only if

$$\exists M, \lambda > 0 \ \forall (x,u,y) \in \mathcal{ZD}(A,B,C) \ \forall t \geqslant 0 \ : \ \|(x(t), u(t))\| \leqslant Me^{-\lambda t}\|x(0)\|.$$

For linear systems (A, B, C) of form (1.1) with ordered vector relative degree $r \in \mathbb{N}^{1\times m}$ the zero dynamics of (A, B, C) can be read off from normal form (2.1) given by Theorem 2.4. Corollary 2.6 provides a characterization of the systems zero dynamics in terms of the normal form. Furthermore, exponential stability of the zero dynamics of (A, B, C) will be characterized.

Corollary 2.6. *For any linear system (A, B, C) of form (1.1) with ordered relative degree $r = (r_1, \ldots, r_m) \in \mathbb{N}^{1\times m}$ and normal form (2.1) and (2.2) the following holds:*

(i) *For $S := [{S^1}^T, \ldots, {S^m}^T]^T$, with $S^1, \ldots, S^m$ defined in (2.16), Γ defined by (2.4), $\mathcal{V}$ defined by (2.7) and Q defined in (2.18), the zero dynamics of (A, B, C) are given by*

$$\mathcal{ZD}(A,B,C) = \Big\{(\mathcal{V}\eta, -\Gamma^{-1}S\eta, 0) \in \mathcal{C}^1([0,\infty) \to \mathbb{R}^n) \times \mathcal{C}^1([0,\infty) \to \mathbb{R}^m) \quad \times \ \mathcal{C}^1([0,\infty) \to \mathbb{R}^m) \Big| \dot{\eta} = Q\eta\Big\}.$$

(ii) *The zero dynamics of (A, B, C) are exponentially stable if, and only if, Q is Hurwitz.*

Given a sufficiently smooth reference signal $y_R \colon \mathbb{R}_{\geqslant 0} \to \mathbb{R}^m$ one can determine an input $u = u_R$ such that the output y of (1.1) matches this signal straightforward by using the normal form (2.1) and (2.2). A system (A, B, C) is called *right-invertible* if this tracking problem can be solved [13]. The following corollary presents the solution to this problem.

Corollary 2.7. *Consider a linear system (A, B, C) of form (1.1) with ordered relative degree $r = (r_1, \ldots, r_m) \in \mathbb{N}^{1\times m}$ and normal form (2.1) and (2.2). Let $y_R = (y_{R1}, \ldots, y_{Rm})^T \colon [0,\infty) \to \mathbb{R}^m$ with $y_{Rj} \in \mathcal{C}^{r_j}([0,\infty) \to \mathbb{R}), j \in \{1, \ldots, m\}$. Let y be the output of (1.1). Then the following are equivalent:*

(i) $y = y_R$,
(ii) *the input u of* (1.1) *is given by*

$$u = u_R = \Gamma^{-1}\left(\begin{bmatrix} {y_{R1}}^{(r_1)} \\ \vdots \\ {y_{Rm}}^{(r_m)} \end{bmatrix} - \begin{bmatrix} R^1 \\ \vdots \\ R^m \end{bmatrix}\xi - \begin{bmatrix} S^1 \\ \vdots \\ S^m \end{bmatrix}\eta\right), \tag{2.19}$$

where for arbitrary $\eta^0 \in \mathbb{R}^{n-r^s}$, η *is a solution of the initial value problem*

$$\dot{\eta} = Q\eta + [P_1, \ldots, P_m]y_R, \quad \eta(0) = \eta^0, \tag{2.20}$$

$\xi = \left(y_{R1}, \ldots, y_{R1}^{(r_1-1)} | y_{R2}, \ldots, y_{R2}^{(r_2-1)} | \cdots | y_{Rm}, \ldots, y_{Rm}^{(r_m-1)}\right)^T$, *Q is defined in* (2.18), $P_1, \ldots, P_m$ *are defined in* (2.17), Γ *is defined in* (2.4), $S^1, \ldots, S^m$ *are defined in* (2.16), $R^j := \left[R^j_{1,1}, \ldots, R^i_{1,r_1} | \cdots | R^j_{m,1}, \ldots, R^i_{m,r_m}\right]$ *and* $R^j_{i,k}$ *is defined in* (2.15).

Systems (2.19) and (2.20) is called the *inverse system* of system (1.1) [9].

3. Proofs

This section contains all proofs for the results given in Section 2. It is structured as follows: First it is shown that for every system with vector relative degree $r \in \mathbb{N}^{1\times m}$ one can find an permutation of the output such that the system with permuted output has an ordered vector relative degree. Next linearly independence of the matrices $\mathcal{C}$ and $\mathcal{B}$, defined by (2.5) and (2.6), respectively, is shown. Then the proof for the normal form including the construction of the coordinate transformation is given. A proof for characterization and stability of the system's zero dynamics is presented and finally right-invertibility of (A, B, C) is shown.

3.1. Ordered vector relative degree

Proof of Lemma 2.3. Let $\sigma : \{1, \ldots, m\} \to \{1, \ldots, m\}$ be a permutation such that $r_{\sigma(1)} \geqslant r_{\sigma(2)} \geqslant \cdots \geqslant r_{\sigma(m)}$. Furthermore, set

$$P := \begin{bmatrix} e^{\sigma(1)}_{(m)} \\ \vdots \\ e^{\sigma(m)}_{(m)} \end{bmatrix}.$$

Then

$$PC = \begin{bmatrix} e^{\sigma(1)}_{(m)} \\ \vdots \\ e^{\sigma(m)}_{(m)} \end{bmatrix}\begin{bmatrix} c^1_{(m)} \\ \vdots \\ c^m_{(m)} \end{bmatrix} = \begin{bmatrix} c^{\sigma(1)}_{(m)} \\ \vdots \\ c^{\sigma(1)}_{(m)} \end{bmatrix}$$

and by the assumption on the relative degree it follows that

$$\forall j \in \{1, \ldots, m\}\ \forall k \in \{0, \ldots, r_{\sigma(j)} - 2\} \ : \ (PC)^j_{(n)}A^kB = c^{\sigma(j)}_{(n)}A^kB = 0_{1\times m}$$

and

$$\operatorname{rk}\begin{bmatrix}(PC)^1_{(n)}A^{r_1-1}B\\ \vdots\\ (PC)^m_{(n)}A^{r_m-1}B\end{bmatrix}=m,$$

whence the linear system (A,B,PC) has relative degree $Pr=(r_{\sigma(1)},\dots,r_{\sigma(m)})$ with $r_{\sigma(1)}\geqslant\cdots\geqslant r_{\sigma(m)}$. $\square$

3.2. Linearly independence of $\mathcal{C}$ and $\mathcal{B}$

Recall the matrices $\mathcal{C}\in\mathbb{R}^{r^s\times n}$, defined by (2.5), and $\mathcal{B}\in\mathbb{R}^{n\times r^s}$ defined by (2.6). Note that, for $m_i, i\in\{1,\dots,r_1\}$, defined by (2.3), it holds true that $m=m_1\geqslant m_2\geqslant\cdots\geqslant m_{r_1}\geqslant 1$ and

$$r^s=\sum_{j=1}^m r_j=\sum_{j=1}^m\sum_{i=1}^{r_1}\frac{\max\{r_j-i+1,0\}}{\max\{r_j-i+1,1\}}=\sum_{i=1}^{r_1}\underbrace{\sum_{j=1}^m\frac{\max\{r_j-i+1,0\}}{\max\{r_j-i+1,1\}}}_{\#\{r_j\,|\,r_j\geqslant i,\,j\in\{1,\dots,m\}\}}=\sum_{i=1}^{r_1}m_i.$$

The following lemma shows that $\mathcal{C}$ and $\mathcal{B}$ have full rank.

Lemma 3.1. *If a linear system (A,B,C) of form (1.1) has ordered vector relative degree $r=(r_1,\dots,r_m)\in\mathbb{N}^{1\times m}$, then $\mathcal{C}$ and $\mathcal{B}$, defined by (2.5) and (2.6), respectively, have full rank.*

Proof. Note that $\sum_{j=1}^m r_j\leqslant n$. It suffices to show that $\mathcal{C}\mathcal{B}\in\mathbb{R}^{r^s\times r^s}$ is invertible.

First consider the first $m_1=m$ rows of $\mathcal{C}\mathcal{B}$. Since (A,B,C) has relative degree $r=(r_1,\dots,r_m)\in N^{1\times m}$ it follows that

$$\mathcal{C}\mathcal{B}=\left[\begin{array}{c}c^1_{(n)}\\ \vdots\\ c^1_{(n)}A^{r_1-1}\\ \hline \vdots\\ \hline c^m_{(n)}\\ \vdots\\ c^m_{(n)}A^{r_m-1}\end{array}\right]B=\left[\begin{array}{c}c^1_{(n)}A^0B\\ \vdots\\ c^1_{(n)}A^{r_1-1}B\\ \hline \vdots\\ \hline c^m_{(n)}A^0B\\ \vdots\\ c^m_{(n)}A^{r_m-1}B\end{array}\right]=\left[\begin{array}{c}0_{(r_1-1)\times m_1}\\ \Gamma^1_{(m_1)}\\ \hline \vdots\\ \hline 0_{(r_m-1)\times m_1}\\ \Gamma^m_{(m_1)}\end{array}\right],\tag{3.1}$$

where $\Gamma^i_{(m_1)}, i\in\{1,\dots,m\}$, is the ith row of Γ. Thus

$$\mathcal{C}\mathcal{B}\Gamma^{-1}[e^{(m)}_1,\dots,e^{(m)}_{m_1}]=\left[\begin{array}{c}0_{(r_1-1)\times m_1}\\ \Gamma^1_{(m_1)}\\ \hline \vdots\\ \hline 0_{(r_m-1)\times m_1}\\ \Gamma^m_{(m_1)}\end{array}\right]\Gamma^{-1}I_{m_1}=\left[\begin{array}{c}0_{(r_1-1)\times m_1}\\ e^1_{(m_1)}\\ \hline \vdots\\ \hline 0_{(r_m-1)\times m_1}\\ e^{m_1}_{(m_1)}\end{array}\right],$$

which shows that $\mathcal{C}\mathcal{B}\Gamma^{-1}[e^{(m)}_1,\dots,e^{(m)}_{m_1}]$ has rank $m_1=m$.

Next consider $\mathcal{C}A^{i-1}B\Gamma^{-1}[e_1^{(m)},\ldots,e_{m_i}^{(m)}]$, for $i \in \{2,\ldots,r_1\}$. Since (A,B,C) has relative degree $r = (r_1,\ldots,r_m) \in \mathbb{N}^{1\times m}$ it follows with the conventions:

(i) $0_{(r_j-i)\times m_i}$ is of dimension zero, if $i \geqslant r_j, j \in \{1,\ldots,m\}$, and
(ii) $\Gamma^j_{(m)}$ and $e^j_{(m_i)}$ do not exist in the following matrices if $j > m_i, j \in \{1,\ldots,m\}$, and
(iii) $\mathcal{X}_{\mu\times\nu} \in \mathbb{R}^{\mu\times\nu}$ is an arbitrarily matrix of dimension $\mu \times \nu$,

that

$$
\begin{aligned}
&\mathcal{C}A^{i-1}B\Gamma^{-1}[e_1^{(m)},\ldots,e_{m_i}^{(m)}]\\
&= \left[\begin{array}{c} c^1_{(n)}A^{i-1}B \\ \vdots \\ c^1_{(n)}A^{r_1-1}B \\ \vdots \\ c^1_{(n)}A^{r_1+i-2}B \\ \hline \vdots \\ \hline c^m_{(n)}A^{i-1}B \\ \vdots \\ c^m_{(n)}A^{r_m-1}B \\ \vdots \\ c^m_{(n)}A^{r_m+i-2}B \end{array}\right] \Gamma^{-1}[e_1^{(m)},\ldots,e_{m_i}^{(m)}]\\
&= \left[\begin{array}{c} 0_{(r_1-i)\times m} \\ \Gamma^1_{(m)} \\ \mathcal{X}_{(\min\{r_1,i-1\})\times m} \\ \hline \vdots \\ \hline 0_{(r_m-i)\times m} \\ \Gamma^m_{(m)} \\ \mathcal{X}_{(\min\{r_m,i-1\})\times m} \end{array}\right] \Gamma^{-1}[e_1^{(m)},\ldots,e_{m_i}^{(m)}]\\
&= \left[\begin{array}{c} 0_{(r_1-i)\times m_i} \\ e^1_{(m_i)} \\ \mathcal{X}_{(\min\{r_1,i-1\})\times m_i} \\ \hline \vdots \\ \hline 0_{(r_m-i)\times m_i} \\ e^m_{(m_i)} \\ \mathcal{X}_{(\min\{r_m,i-1\})\times m_i} \end{array}\right], \quad i \in \{2,\ldots,r_1\}.
\end{aligned}
$$

Thus, for all $i \in \{1, \dots, r_1\}$, the m_i rows of $\mathcal{C}A^{i-1}B\Gamma^{-1}[e_1^{(m)}, \dots, e_{m_i}^{(m)}]$ are linearly independent, and since

$$\mathcal{CB} = \begin{array}{c} \left[\begin{array}{c|c|c|c|c} \begin{matrix} 0_{(r_1-1)\times m_1} \\ e^1_{(m_1)} \end{matrix} & \begin{matrix} 0_{(r_1-2)\times m_2} \\ e^1_{(m_2)} \\ \mathcal{X}_{1\times m_2} \end{matrix} & \begin{matrix} \iddots \\ \iddots \\ \iddots \end{matrix} & \begin{matrix} 0_{1\times m_{r_1-1}} \\ e^1_{(m_{r_1-1})} \\ \mathcal{X}_{(r_1-2)\times m_{r_1-1}} \end{matrix} & \begin{matrix} e^1_{(m_{r_1})} \\ \mathcal{X}_{(r_1-1)\times m_{r_1}} \end{matrix} \\ \hline \begin{matrix} 0_{(r_2-1)\times m_1} \\ e^2_{(m_1)} \end{matrix} & \begin{matrix} 0_{(r_2-2)\times m_2} \\ e^2_{(m_2)} \\ \mathcal{X}_{1\times m_2} \end{matrix} & \begin{matrix} \iddots \\ \iddots \\ \iddots \end{matrix} & \begin{matrix} 0_{(r_2-r_1+1)\times m_{r_1-1}} \\ e^2_{(m_{r_1-1})} \\ \mathcal{X}_{(\min\{r_1-2,r_2\})\times m_{r_1-1}} \end{matrix} & \begin{matrix} e^2_{(m_{r_1})} \\ \mathcal{X}_{(\min\{r_1-1,r_2\})\times m_{r_1}} \end{matrix} \\ \hline \vdots & \vdots & & \vdots & \vdots \\ \hline \begin{matrix} 0_{(r_m-1)\times m_1} \\ e^m_{(m_1)} \end{matrix} & \begin{matrix} 0_{(r_m-2)\times m_2} \\ e^m_{(m_2)} \\ \mathcal{X}_{1\times m_2} \end{matrix} & \begin{matrix} \iddots \\ \iddots \\ \iddots \end{matrix} & \begin{matrix} 0_{(r_m-r_1+1)\times m_{r_1-1}} \\ e^m_{(m_{r_1-1})} \\ \mathcal{X}_{(\min\{r_1-2,r_m\})\times m_{r_1-1}} \end{matrix} & \begin{matrix} e^m_{(m_{r_1})} \\ \mathcal{X}_{(\min\{r_1-1,r_m\})\times m_{r_1}} \end{matrix} \end{array}\right] \begin{matrix} \}\, r_1 \\ \}\, r_2 \\ \vdots \\ \}\, r_m \end{matrix} \\ \begin{matrix} \underbrace{\qquad}_{m_1} & \underbrace{\qquad}_{m_2} & \cdots & \underbrace{\qquad}_{m_{r_1-1}} & \underbrace{\qquad}_{m_{r_1}} \end{matrix} \end{array} \tag{3.2}$$

it follows that $\mathcal{CB}$ is invertible. $\square$

As an immediate consequence of Lemma 3.1 it follows that for linear systems (A, B, C) of form (1.1) with vector relative degree $r = (r_1, \dots, r_m) \in \mathbb{N}^{1\times m}$, the matrices $C \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$ have full rank m.

3.3. Coordinate transformation and normal form

Lemma 3.1 shows that the rows of $\mathcal{C}$ qualify as basis, which, if $r^s = \sum_{j=1}^m r_j < n$, has to be completed, for a coordinate transformation in $\mathbb{R}^n$. Consider a matrix $\mathcal{V} \in \mathbb{R}^{n\times(n-r^s)}$, given by (2.7).

For $\widehat{U}$ and $\mathcal{N}$, given by (2.8), it follows from

$$\begin{bmatrix} \mathcal{C} \\ \mathcal{N} \end{bmatrix} [\mathcal{B}(\mathcal{CB})^{-1}, \mathcal{V}] = I_n,$$

that $\widehat{U}$ has the inverse

$$\widehat{U}^{-1} = [\mathcal{B}(\mathcal{CB})^{-1}, \mathcal{V}]. \tag{3.3}$$

Although $\widehat{U}$ already qualifies as coordinate transformation in $\mathbb{R}^n$ we do not obtain a normal form which has the same structure properties as the normal form (1.4) for linear SISO-systems (1.3), i.e. the transformation matrix $\widehat{U}$ will not lead in general to a matrix $\widetilde{A}$ as in (2.2). Therefore, it is necessary to consider the transformation matrix U, given by (2.14) and $T_i, \widehat{\mathcal{C}}_i, \widehat{\mathcal{B}}_i, \widehat{\mathcal{N}}_i, \widehat{U}_i$, for $i \in \{1, \dots, m\}$, defined in (2.9)–(2.13), respectively.

Proof of Theorem 2.4. *Step 1*: First it is shown that the coordinate transformation

$$\begin{pmatrix} \chi \\ \zeta \end{pmatrix} := \widehat{U}x, \quad \chi(t) = (y_1(t), \dots, y_1^{(r_1-1)}(t)| \cdots |y_m(t), \dots, y_m^{(r_m-1)}(t))^T \in \mathbb{R}^{r^s},$$
$$\zeta(t) \in \mathbb{R}^{n-r^s},$$

given by (2.7) and (2.8), converts (1.1) into

$$\left.\begin{aligned} &\tfrac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}\chi\\ \zeta\end{pmatrix} = \widehat{A}\begin{pmatrix}\chi\\ \zeta\end{pmatrix} + \widetilde{B}u \\ &y = \widetilde{C}\begin{pmatrix}\chi\\ \zeta\end{pmatrix}\end{aligned}\right\}, \tag{3.4}$$

where

$$\widehat{A} = \left[\begin{array}{cccc|c} \widehat{A}_{1,1} & \widehat{A}_{1,2} & \cdots & \widehat{A}_{1,m} & \widehat{S}_1 \\ \widehat{A}_{2,1} & \widehat{A}_{2,2} & \cdots & \widehat{A}_{2,m} & \widehat{S}_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \widehat{A}_{m,1} & \widehat{A}_{m,2} & \cdots & \widehat{A}_{m,m} & \widehat{S}_m \\ \hline \widehat{P}_1 & \widehat{P}_2 & \cdots & \widehat{P}_m & \widehat{Q} \end{array}\right] \tag{3.5}$$

and for $i, j \in \{1, \ldots, m\}$

$$\begin{aligned} \widehat{A}_{i,i} &:= \begin{bmatrix} 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & 1 \\ \widehat{R}^i_{i,1} & \cdots & & \widehat{R}^i_{i,r_i} \end{bmatrix} \in \mathbb{R}^{r_i \times r_i}, \\ \widehat{A}_{i,j} &:= \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \\ \widehat{R}^i_{j,1} & \cdots & \widehat{R}^i_{j,r_j} \end{bmatrix} \in \mathbb{R}^{r_i \times r_j}, \quad j \neq i, \end{aligned} \tag{3.6}$$

where $\widehat{R}^j_{i,k} \in \mathbb{R}$, for $k \in \{1, \ldots, r_i\}$ and $i, j \in \{1, \ldots, m\}$, and, for $i \in \{1, \ldots, m\}$

$$\widehat{S}_i := \begin{bmatrix} 0_{(r_i-1)\times(n-r^s)} \\ S^i \end{bmatrix} \in \mathbb{R}^{r_i \times (n-r^s)}, \quad \widehat{P}_i \in \mathbb{R}^{(n-r^s)\times r_i}, \quad \widehat{Q} \in \mathbb{R}^{(n-r^s)\times(n-r^s)}. \tag{3.7}$$

Step 1a: First the structure of $\widehat{A}$ is proven. By definition of $\widehat{U}$, see (2.8), it follows that

$$\widehat{A} = \widehat{U}A\widehat{U}^{-1} = \begin{bmatrix} \mathcal{C} \\ \mathcal{N} \end{bmatrix} A[\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, \mathcal{V}] = \left[\begin{array}{c|c} \mathcal{C}A\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} & \mathcal{C}A\mathcal{V} \\ \hline \mathcal{N}A\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} & \mathcal{N}A\mathcal{V} \end{array}\right].$$

Thus

$$[\widehat{P}_1, \ldots, \widehat{P}_m] = \mathcal{N}A\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} \in \mathbb{R}^{(n-r^s)\times r^s}, \tag{3.8}$$

$$\widehat{Q} = \mathcal{N}A\mathcal{V} \in \mathbb{R}^{(n-r^s)\times(n-r^s)} \tag{3.9}$$

and the definition of $\mathcal{C}$ and $\mathcal{B}$, see (2.5) and (2.6), respectively, yields

$$
\begin{aligned}
\mathcal{C}A\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} &= \left[\begin{array}{c} c^1_{(n)}A \\ \vdots \\ c^1_{(n)}A^{r_1} \\ \hline \vdots \\ \hline c^m_{(n)}A \\ \vdots \\ c^m_{(n)}A^{r_m} \end{array}\right] \mathcal{B}(\mathcal{C}\mathcal{B})^{-1} = \left[\begin{array}{c} \mathcal{C}^2_{(n)} \\ \vdots \\ \mathcal{C}^{r_1}_{(n)} \\ c^1_{(n)}A^{r_1} \\ \hline \vdots \\ \hline \mathcal{C}^{\sum_{j=1}^{m-1} r_j+2}_{(n)} \\ \vdots \\ \mathcal{C}^{r^s}_{(n)} \\ c^m_{(n)}A^{r_m} \end{array}\right] \mathcal{B}(\mathcal{C}\mathcal{B})^{-1} = \left[\begin{array}{c} (\mathcal{C}\mathcal{B})^2_{(r^s)} \\ \vdots \\ (\mathcal{C}\mathcal{B})^{r_1}_{(r^s)} \\ c^1_{(n)}A^{r_1}\mathcal{B} \\ \hline \vdots \\ \hline (\mathcal{C}\mathcal{B})^{\sum_{j=1}^{m-1} r_j+2}_{(r^s)} \\ \vdots \\ (\mathcal{C}\mathcal{B})^{r^s}_{(r^s)} \\ c^m_{(n)}A^{r_m}\mathcal{B} \end{array}\right] (\mathcal{C}\mathcal{B})^{-1} \\
&= \left[\begin{array}{c} e^2_{(r^s)} \\ \vdots \\ e^{r_1}_{(r^s)} \\ c^1_{(n)}A^{r_1}\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} \\ \hline \vdots \\ \hline e^{\sum_{j=1}^{m-1} r_j+2}_{(r^s)} \\ \vdots \\ e^{r^s}_{(r^s)} \\ c^m_{(n)}A^{r_m}\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} \end{array}\right].
\end{aligned} \tag{3.10}
$$

Furthermore, invoking $\operatorname{im}\mathcal{V} = \ker\mathcal{C}$, it follows that

$$
\mathcal{C}A\mathcal{V} = \left[\begin{array}{c} c^1_{(n)}A \\ \vdots \\ c^1_{(n)}A^{r_1-1} \\ c^1_{(n)}A^{r_1} \\ \hline \vdots \\ \hline c^m_{(n)}A \\ \vdots \\ c^m_{(n)}A^{r_m-1} \\ c^m_{(n)}A^{r_m} \end{array}\right] \mathcal{V} = \left[\begin{array}{c} \mathcal{C}^2_{(n)} \\ \vdots \\ \mathcal{C}^{r_1}_{(n)} \\ c^1_{(n)}A^{r_1} \\ \hline \vdots \\ \hline (\mathcal{C})^{\sum_{j=1}^{m-1} r_j+2}_{(n)} \\ \vdots \\ (\mathcal{C})^{r^s}_{(n)} \\ c^m_{(n)}A^{r_m} \end{array}\right] \mathcal{V} = \left[\begin{array}{c} 0_{1\times(n-r^s)} \\ \vdots \\ 0_{1\times(n-r^s)} \\ c^1_{(n)}A^{r_1}\mathcal{V} \\ \hline \vdots \\ \hline 0_{1\times(n-r^s)} \\ \vdots \\ 0_{1\times(n-r^s)} \\ c^m_{(n)}A^{r_m}\mathcal{V} \end{array}\right] \overset{(2.16)}{=} \left[\begin{array}{c} 0_{(r_1-1)\times(n-r^s)} \\ S^1 \\ \hline \vdots \\ \hline 0_{(r_m-1)\times(n-r^s)} \\ S^m \end{array}\right]. \tag{3.11}
$$

Hence, setting

$$
\left[\widehat{R}^i_{1,1},\ldots,\widehat{R}^i_{1,r_1}\middle|\cdots\middle|\widehat{R}^i_{m,1},\ldots,\widehat{R}^i_{m,r_m}\right] := c^i_{(n)}A^{r_i}\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, \quad i\in\{1,\ldots,m\}, \tag{3.12}
$$

(3.10) and (3.11) yield the structure of $\widehat{A}$ as given in (3.5)–(3.7).

Step 1b: Next the structure of $\widetilde{B}$ is proven. By the definition of $\widehat{U}$, see (2.8), it follows that

$$\widetilde{B} = \widehat{U}B = \begin{bmatrix} \mathcal{C}B \\ (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T[B - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}B] \end{bmatrix}.$$

Recall (3.1), i.e.

$$\mathcal{C}B = \begin{bmatrix} 0_{(r_1-1)\times m} \\ \Gamma^1_{(m)} \\ \hline \vdots \\ \hline 0_{(r_m-1)\times m} \\ \Gamma^m_{(m)} \end{bmatrix} = \begin{bmatrix} 0_{(r_1-1)\times m} \\ c^1_{(n)}A^{r_1-1}B \\ \hline \vdots \\ \hline 0_{(r_m-1)\times m} \\ c^m_{(n)}A^{r_m-1}B \end{bmatrix}.$$

Furthermore

$$\begin{aligned} &(\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T[B - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}B] \\ &= (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[B\Gamma^{-1}\Gamma - \mathcal{B}(\mathcal{C}\mathcal{B})^{-1}\mathcal{C}\underbrace{B\Gamma^{-1}[e_1^{(m)},\ldots,e_m^{(m)}]}_{=[\mathcal{B}_1^{(n)},\ldots,\mathcal{B}_m^{(n)}]}\Gamma\right] \\ &= (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T\left[B\Gamma^{-1}\Gamma - \mathcal{B}\underbrace{(\mathcal{C}\mathcal{B})^{-1}[(\mathcal{C}\mathcal{B})_1^{(r^s)},\ldots,(\mathcal{C}\mathcal{B})_m^{(r^s)}]}_{=[e_1^{(r^s)},\ldots,e_m^{(r^s)}]}\Gamma\right] \\ &= (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T([[\mathcal{B}_1^{(n)},\ldots,\mathcal{B}_m^{(n)}] - [\mathcal{B}_1^{(n)},\ldots,\mathcal{B}_m^{(n)}])\Gamma \\ &= 0_{(n-r^s)\times m}, \end{aligned}$$

which shows the structure of $\widetilde{B}$ as in (2.2).

Step 1c: Now the structure of $\widetilde{C}$ is shown. Since the rows of C are also rows of $\mathcal{C}$, i.e.

$$C = \begin{bmatrix} c^1_{(n)} \\ c^2_{(n)} \\ \vdots \\ c^m_{(n)} \end{bmatrix} = \begin{bmatrix} \mathcal{C}^1_{(n)} \\ \mathcal{C}^{r_1+1}_{(n)} \\ \vdots \\ \mathcal{C}^{r^s-r_m+1}_{(n)} \end{bmatrix}$$

and since $\operatorname{im}\mathcal{V} = \ker\mathcal{C}$ it follows that $C\mathcal{V} = 0_{m\times(n-r^s)}$. Furthermore

$$\begin{aligned} C\mathcal{B}(\mathcal{C}\mathcal{B})^{-1} &= \begin{bmatrix} (\mathcal{C}\mathcal{B})^1_{(r^s)} \\ (\mathcal{C}\mathcal{B})^{r_1+1}_{(r^s)} \\ \vdots \\ (\mathcal{C}\mathcal{B})^{r^s-r_m+1}_{(r^s)} \end{bmatrix}(\mathcal{C}\mathcal{B})^{-1} \\ &= \begin{bmatrix} 1 & 0_{1\times(r_1-1)} & 0 & 0_{1\times(r_2-1)} & 0 & \cdots & 0 & 0_{1\times(r_m-1)} \\ 0 & 0_{1\times(r_1-1)} & 1 & 0_{1\times(r_2-1)} & 0 & \cdots & 0 & 0_{1\times(r_m-1)} \\ 0 & 0_{1\times(r_1-1)} & 0 & 0_{1\times(r_2-1)} & 1 & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & 0 & 0_{1\times(r_m-1)} \\ 0 & 0_{1\times(r_1-1)} & 0 & 0_{1\times(r_2-1)} & 0 & & 1 & 0_{1\times(r_m-1)} \end{bmatrix}. \end{aligned}$$

Hence

$$\widetilde{C} = C\widehat{U}^{-1} = [C\mathcal{B}(\mathcal{C}\mathcal{B})^{-1}, C\mathcal{V}]$$

yields the structure of $\widetilde{C}$ as in (2.2).

Step 2: We show that the coordinate transformation

$$\begin{pmatrix}\xi\\ \eta\end{pmatrix} := Ux, \quad \xi(t) = \left(y_1(t), \ldots, y_1^{(r_1-1)}(t)\middle| \cdots \middle| y_m(t), \ldots, y_m^{(r_m-1)}(t)\right)^T \in \mathbb{R}^{r^s},$$
$$\eta(t) \in \mathbb{R}^{n-r^s}$$

given by (2.7)–(2.14) converts the linear system (A, B, C) of form (1.1) into (2.1) with $\widetilde{A}, \widetilde{B}, \widetilde{C}$ as in (2.2) with matrix components of $\widetilde{A}$ as in (2.15)–(2.18).

Recall the structure of $\widehat{A}$ given by (3.5)–(3.7). For $i \in \{1, \ldots, m\}$, consider the matrices

$$\widehat{A}_i := \left[\begin{array}{cccc|c} 0 & 1 & \cdots & 0 & \\ \vdots & & \ddots & & \\ 0 & \cdots & 0 & 1 & \widehat{S}_i \\ \widehat{R}^i_{i,1} & \cdots & & \widehat{R}^i_{i,r_i} & \\ \hline & & \widehat{P}_i & & \widehat{Q} \end{array}\right] = \left[\begin{array}{c|c} \widehat{A}_{i,i} & \widehat{S}_i \\ \hline \widehat{P}_i & \widehat{Q} \end{array}\right] = T_i \widehat{A} T_i^T$$
$$= T_i \widehat{U} A \widehat{U}^{-1} T_i^T \in \mathbb{R}^{(r_i+n-r^s)\times(r_i+n-r^s)},$$
$$\widehat{C}_i := [1, 0_{1\times(r_i+n-r^s-1)}] = e^1_{(r_i+n-r^s)} = e^i_{(m)} \widetilde{C} T_i^T \in \mathbb{R}^{1\times(r_i+n-r^s)},$$
$$\widehat{B}_i := \left[\begin{array}{c} 0_{(r_i-1)\times 1} \\ 1 \\ \hline 0_{(n-r^s)\times 1} \end{array}\right] = e^{(r_i+n-r^s)}_{r_i} = T_i \widetilde{B} \Gamma^{-1} e^{(m)}_i \in \mathbb{R}^{r_i+n-r^s}.$$

Then

$$\begin{bmatrix} \widehat{C}_i \\ \widehat{C}_i\widehat{A}_i \\ \vdots \\ \widehat{C}_i\widehat{A}_i^{r_i-1} \end{bmatrix} = \begin{bmatrix} e^1_{(r_i+n-r^s)} \\ e^2_{(r_i+n-r^s)} \\ \vdots \\ e^{r_i}_{(r_i+n-r^s)} \end{bmatrix} = [I_{r_i}, 0_{r_i\times(n-r^s)}] \stackrel{(2.10)}{=} \widehat{\mathcal{C}}_i \in \mathbb{R}^{r_i\times(r_i+n-r^s)},$$
$$[\widehat{B}_i, \widehat{A}_i\widehat{B}_i, \ldots, \widehat{A}_i^{r_i-1}\widehat{B}_i] = [e^{(r_i+n-r^s)}_{r_i}, \ldots, \widehat{A}_i^{r_i-1} e^{(r_i+n-r^s)}_{r_i}] \stackrel{(2.11)}{=} \widehat{\mathcal{B}}_i \in \mathbb{R}^{(r_i+n-r^s)\times r_i}.$$

More precisely $\widehat{\mathcal{B}}_i$ is structured as follows:

$$\widehat{\mathcal{B}}_i = \left[\begin{array}{c|cccc} 0 & 0 & \cdots & 0 & 1 \\ \vdots & \vdots & \iddots & 1 & * \\ 0 & 0 & \iddots & \iddots & \vdots \\ 0 & 1 & * & \cdots & * \\ 1 & * & \cdots & & * \\ \hline 0_{(n-r^s)\times 1} & & \mathcal{X}_{(n-r^s)\times(r_i-1)} & & \end{array}\right] \in \mathbb{R}^{(r_i+n-r^s)\times r_i}. \tag{3.13}$$

Since $\widehat{C}_i\widehat{A}_i^j\widehat{B}_i = 0$, for all $j \in \{0, \ldots, r_i - 2\}$, and $\widehat{C}_i\widehat{A}_i^{r_i-1}\widehat{B}_i = 1$, it follows that the linear SISO-system:

$$\left.\begin{array}{l} \dot{z} = \widehat{A}_i z + \widehat{B}_i v \\ w = \widehat{C}_i z \end{array}\right\}$$

has relative degree r_i. Furthermore, it follows that

$$\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i = \begin{bmatrix} 0 & \cdots & 0 & 1 \\ \vdots & ⋰ & ⋰ & * \\ 0 & 1 & ⋰ & \vdots \\ 1 & * & \cdots & * \end{bmatrix} \quad \text{and} \quad (\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1} = \begin{bmatrix} * & \cdots & * & 1 \\ \vdots & ⋰ & ⋰ & 0 \\ * & 1 & ⋰ & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix} \tag{3.14}$$

and thus

$$\widehat{B}_i\,(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1} = \left[\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ * & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ * & \cdots & * & 1 & 0 \\ * & \cdots & & * & 1 \\ \hline \multicolumn{4}{c|}{\mathcal{X}_{(n-r^s)\times(r_i-1)}} & 0_{(n-r^s)\times 1} \end{array}\right] \in \mathbb{R}^{(r_i+n-r^s)\times r_i}, \tag{3.15}$$

$$[0_{(n-r^s)\times r_i}, I_{n-r^s}]\widehat{B}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1} = [\mathcal{X}_{(n-r^s)\times(r_i-1)}, 0_{(n-r^s)\times 1}]. \tag{3.16}$$

Set $\widehat{\mathcal{V}}_i := \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix}$. Then $\ker\widehat{\mathcal{C}}_i = \operatorname{im}\widehat{\mathcal{V}}_i$ and thus

$$\begin{bmatrix} [I_{r_i}, 0_{r_i\times(n-r^s)}] \\ \underbrace{(\widehat{\mathcal{V}}_i\widehat{\mathcal{V}}_i^T)^{-1}\widehat{\mathcal{V}}_i^T}_{=0[0_{(n-r^s)\times r_i}, I_{n-r^s}]} [I_{r_i+n-r^s} - \widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1}\widehat{\mathcal{C}}_i] \end{bmatrix} \overset{(2.10),(2.12)}{=} \begin{bmatrix} \widehat{\mathcal{C}}_i \\ \widehat{\mathcal{N}}_i \end{bmatrix} \tag{3.17}$$

is invertible with inverse

$$\begin{bmatrix} \widehat{\mathcal{C}}_i \\ \widehat{\mathcal{N}}_i \end{bmatrix}^{-1} = [\widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1}, \widehat{\mathcal{V}}_i]. \tag{3.18}$$

Furthermore

$$\begin{aligned} \widehat{\mathcal{C}}_i \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix} &= [I_{r_i}, 0_{r_i\times(n-r^s)}] \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix} = 0_{r_i\times(n-r^s)}, \\ \widehat{\mathcal{C}}_i \begin{bmatrix} I_{r_i} \\ 0_{(n-r^s)\times r_i} \end{bmatrix} &= [I_{r_i}, 0_{r_i\times(n-r^s)}] \begin{bmatrix} I_{r_i} \\ 0_{(n-r^s)\times r_i} \end{bmatrix} = I_{r_i}, \\ \widehat{\mathcal{N}}_i \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix} &= [0_{(n-r^s)\times r_i}, I_{n-r^s}] \left[\begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix} - \widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1}\widehat{\mathcal{C}}_i \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix} \right] = I_{n-r^s} \end{aligned}$$

and thus

$$\left[\begin{array}{c|c|c|c} I_{\sum_{j=1}^{i-1} r_j} & 0_{\left(\sum_{j=1}^{i-1} r_j\right)\times r_i} & 0_{\left(\sum_{j=1}^{i-1} r_j\right)\times\left(\sum_{j=i+1}^{m} r_j\right)} & 0_{\left(\sum_{j=1}^{i-1} r_j\right)\times(n-r^s)} \\ \hline 0_{r_i\times\left(\sum_{j=1}^{i-1} r_j\right)} & \underbrace{\widehat{\mathcal{C}}_i \begin{bmatrix} I_{r_i} \\ 0_{(n-r^s)\times r_i} \end{bmatrix}}_{=I_{r_i}} & 0_{r_i\times\left(\sum_{j=1}^{i-1} r_j\right)} & \underbrace{\widehat{\mathcal{C}}_i \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix}}_{=0_{r_i\times(n-r^s)}} \\ \hline 0_{\left(\sum_{j=i+1}^{m} r_j\right)\times\left(\sum_{j=1}^{i-1} r_j\right)} & 0_{\left(\sum_{j=i+1}^{m} r_j\right)\times r_i} & I_{\sum_{j=i+1}^{m} r_j} & 0_{\left(\sum_{j=i+1}^{m} r_j\right)\times(n-r^s)} \\ \hline 0_{(n-r^s)\times\left(\sum_{j=1}^{i-1} r_j\right)} & \widehat{\mathcal{N}}_i \begin{bmatrix} I_{r_i} \\ 0_{(n-r^s)\times r_i} \end{bmatrix} & 0_{(n-r^s)\times\left(\sum_{j=i+1}^{m} r_j\right)} & \underbrace{\widehat{\mathcal{N}}_i \begin{bmatrix} 0_{r_i\times(n-r^s)} \\ I_{n-r^s} \end{bmatrix}}_{=I_{n-r^s}} \end{array}\right] \overset{(2.13)}{=} \widehat{U}_i$$

and, since

$$[I_{r_i}, 0_{r_i\times(n-r^s)}]\widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1} = \widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1} = I_{r_i},$$

it follows from (3.17) and (3.18) that

$$\widehat{U}_i^{-1} = \left[\begin{array}{c|c|c|c} I_{\sum_{j=1}^{i-1} r_j} & 0_{\left(\sum_{j=1}^{i-1} r_j\right)\times r_i} & 0_{\left(\sum_{j=1}^{i-1} r_j\right)\times\left(\sum_{j=i+1}^{m} r_j\right)} & 0_{\left(\sum_{j=1}^{i-1} r_j\right)\times(n-r^s)} \\ \hline 0_{r_i\times\left(\sum_{j=1}^{i-1} r_j\right)} & \underbrace{\begin{array}{c}[I_{r_i}, 0_{r_i\times(n-r^s)}] \\ \cdot\widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1}\end{array}}_{=I_{r_i}} & 0_{r_i\times\left(\sum_{j=1}^{i-1} r_j\right)} & \underbrace{[I_{r_i}, 0_{r_i\times(n-r^s)}]\widehat{\mathcal{V}}_i}_{=0_{r_i\times(n-r^s)}} \\ \hline 0_{\left(\sum_{j=i+1}^{m} r_j\right)\times\left(\sum_{j=1}^{i-1} r_j\right)} & 0_{\left(\sum_{j=i+1}^{m} r_j\right)\times r_i} & I_{\sum_{j=i+1}^{m} r_j} & 0_{\left(\sum_{j=i+1}^{m} r_j\right)\times(n-r^s)} \\ \hline 0_{(n-r^s)\times\left(\sum_{j=1}^{i-1} r_j\right)} & \begin{array}{c}[0_{(n-r^s)\times r_i}, I_{n-r^s}] \\ \cdot\widehat{\mathcal{B}}_i(\widehat{\mathcal{C}}_i\widehat{\mathcal{B}}_i)^{-1}\end{array} & 0_{(n-r^s)\times\left(\sum_{j=i+1}^{m} r_j\right)} & \underbrace{[0_{(n-r^s)\times r_i}, I_{n-r^s}]\widehat{\mathcal{V}}_i}_{=I_{n-r^s}} \end{array}\right]. \tag{3.19}$$

Recall $\widehat{A} = \widehat{U}A\widehat{U}^{-1}$ given by (3.5)–(3.7). First apply the transformation $\widehat{U}_1$. Then, omitting the dimensions of the zeros and identity matrices in $\widehat{U}_1$, it follows that

$$\widehat{U}_1\widehat{A}\widehat{U}_1^{-1} = \left[\begin{array}{c|c|c} I & 0 & 0 \\ \hline 0 & I & 0 \\ \hline \widehat{\mathcal{N}}_1\begin{bmatrix} I \\ 0\end{bmatrix} & 0 & \underbrace{\widehat{\mathcal{N}}_1\begin{bmatrix} 0 \\ I\end{bmatrix}}_{=I} \end{array}\right]\left[\begin{array}{c|c|c} \widehat{A}_{1,1} & \ldots\quad \widehat{A}_{1,m} & \widehat{S}_1 \\ \hline \vdots & \ddots\quad \vdots & \vdots \\ \widehat{A}_{m,1} & \ldots\quad \widehat{A}_{m,m} & \widehat{S}_m \\ \hline \widehat{P}_1 & \ldots\quad \widehat{P}_m & \widehat{Q} \end{array}\right]$$

$$\times\left[\begin{array}{c|c|c} \underbrace{[I,0]\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}}_{=I} & 0 & 0 \\ \hline 0 & I & 0 \\ [0,I]\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1} & 0 & I \end{array}\right]$$

$$= \left[\begin{array}{c|c|c} \widehat{A}_{1,1} & \widehat{A}_{1,2}\quad \ldots\quad \widehat{A}_{1,m} & \widehat{S}_1 \\ \hline \widehat{A}_{2,1} & \widehat{A}_{2,2}\quad \ldots\quad \widehat{A}_{2,m} & \widehat{S}_2 \\ \vdots & \vdots \qquad\qquad \vdots & \vdots \\ \widehat{A}_{m,1} & \widehat{A}_{m,2}\quad \ldots\quad \widehat{A}_{m,m} & \widehat{S}_m \\ \hline \widehat{\mathcal{N}}_1\begin{bmatrix} I \\ 0\end{bmatrix}\widehat{A}_{1,1} + \widehat{\mathcal{N}}_1\begin{bmatrix} 0 \\ I\end{bmatrix}\widehat{P}_1 & \begin{array}{c}\widehat{\mathcal{N}}_1\begin{bmatrix} I \\ 0\end{bmatrix}[\widehat{A}_{1,2},\ldots,\widehat{A}_{1,m}] \\ +I[\widehat{P}_2,\ldots,\widehat{P}_m]\end{array} & \widehat{\mathcal{N}}_1\begin{bmatrix} I \\ 0\end{bmatrix}\widehat{S}_1 + \widehat{\mathcal{N}}_1\begin{bmatrix} 0 \\ I\end{bmatrix}\widehat{Q} \end{array}\right]$$

$$\cdot\left[\begin{array}{c|c|c} \underbrace{[I,0]\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}}_{=I} & 0 & 0 \\ \hline 0 & I & 0 \\ \hline [0,I]\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1} & 0 & I \end{array}\right]$$

$$
= \left[\begin{array}{c|ccc|c}
\widehat{A}_{1,1} I + \widehat{S}_1 [0, I] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} & \widehat{A}_{1,2} & \dots & \widehat{A}_{1,m} & \widehat{S}_1 \\
\hline
\begin{bmatrix} \widehat{A}_{2,1} \\ \vdots \\ \widehat{A}_{m,1} \end{bmatrix} + \begin{bmatrix} \widehat{S}_2 \\ \vdots \\ \widehat{S}_m \end{bmatrix} [0, I] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} & \begin{matrix} \widehat{A}_{2,2} \\ \vdots \\ \widehat{A}_{m,2} \end{matrix} & \begin{matrix} \dots \\ \\ \dots \end{matrix} & \begin{matrix} \widehat{A}_{2,m} \\ \vdots \\ \widehat{A}_{m,m} \end{matrix} & \begin{matrix} \widehat{S}_2 \\ \vdots \\ \widehat{S}_m \end{matrix} \\
\hline
\begin{matrix} \widehat{\mathcal{N}}_1 \left[\begin{bmatrix} I \\ 0 \end{bmatrix} \widehat{A}_{1,1} + \begin{bmatrix} 0 \\ I \end{bmatrix} \widehat{P}_1 \right] [I, 0] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \\ + \widehat{\mathcal{N}}_1 \left[\begin{bmatrix} I \\ 0 \end{bmatrix} \widehat{S}_1 + \begin{bmatrix} 0 \\ I \end{bmatrix} \widehat{Q} \right] [0, I] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \end{matrix} & \multicolumn{3}{c|}{\begin{matrix} \widehat{\mathcal{N}}_1 \begin{bmatrix} I \\ 0 \end{bmatrix} [\widehat{A}_{1,2}, \dots, \widehat{A}_{1,m}] \\ + [\widehat{P}_2, \dots, \widehat{P}_m] \end{matrix}} & \widehat{\mathcal{N}}_1 \begin{bmatrix} I \\ 0 \end{bmatrix} \widehat{S}_1 + \widehat{Q}
\end{array}\right]. \tag{3.20}
$$

Furthermore, for $j \in \{2, \dots, m\}$

$$
\begin{aligned}
&\widehat{A}_{1,1} I_{r_1} + \widehat{S}_1 [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \\
&\quad = \widehat{A}_{1,1} + \begin{bmatrix} 0_{(r_1-1) \times (n-r^s)} \\ S^1 \end{bmatrix} [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \\
&\quad \overset{(3.11)}{=} \begin{bmatrix} 0_{(r_1-1) \times 1}, I_{r_1-1} \\ [\widehat{R}^i_{1,1}, \dots, \widehat{R}^i_{1,r_1}] \end{bmatrix} + \begin{bmatrix} 0_{(r_1-1) \times (n-r^s)} \\ c^1_{(n)} A^{r_1} \mathcal{V} \end{bmatrix} [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \\
&\quad \overset{(3.12)}{=} \begin{bmatrix} 0_{(r_1-1) \times 1}, I_{r_1-1} \\ c^1_{(n)} A^{r_1} \mathcal{B} (\mathcal{C}\mathcal{B})^{-1} \begin{bmatrix} I_{r_1} \\ 0_{(r_s-r_1) \times r_1} \end{bmatrix} \end{bmatrix} + \begin{bmatrix} 0_{(r_1-1) \times (n-r^s)} \\ c^1_{(n)} A^{r_1} \mathcal{V} \end{bmatrix} [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1},
\end{aligned} \tag{3.21}
$$

$$
\begin{aligned}
&\widehat{A}_{j,1} + \widehat{S}_j [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \\
&\quad \overset{(3.11)}{=} \begin{bmatrix} 0_{(r_j-1) \times r_1} \\ [\widehat{R}^j_{1,1}, \dots, \widehat{R}^j_{1,r_1}] \end{bmatrix} + \begin{bmatrix} 0_{(r_j-1) \times (n-r^s)} \\ c^j_{(n)} A^{r_j} \mathcal{V} \end{bmatrix} [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \\
&\quad \overset{(3.12)}{=} \begin{bmatrix} 0_{(r_j-1) \times r_1} \\ c^j_{(n)} A^{r_j} \mathcal{B} (\mathcal{C}\mathcal{B})^{-1} \begin{bmatrix} I_{r_1} \\ 0_{(r_s-r_1) \times r_1} \end{bmatrix} \end{bmatrix} + \begin{bmatrix} 0_{(r_j-1) \times (n-r^s)} \\ c^j_{(n)} A^{r_j} \mathcal{V} \end{bmatrix} [0_{(n-r^s) \times r_1}, I_{r_1}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1},
\end{aligned} \tag{3.22}
$$

$$
\begin{aligned}
&\widehat{\mathcal{N}}_1 \begin{bmatrix} I_{r_1} \\ 0_{(n-r^s) \times r_1} \end{bmatrix} \widehat{A}_{1,j} \\
&\quad \overset{(3.17)}{=} [0_{(n-r^s) \times r_1}, I_{n-r^s}] [I_{r_1+n-r^s} - \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \widehat{\mathcal{C}}_1] \begin{bmatrix} I_{r_1} \\ 0_{(n-r^s) \times r_1} \end{bmatrix} \begin{bmatrix} 0_{(r_1-1) \times r_j} \\ [\widehat{R}^1_{j,1}, \dots, \widehat{R}^1_{j,r_j}] \end{bmatrix} \\
&\quad \overset{(2.10)}{=} [0_{(n-r^s) \times r_1}, I_{n-r^s}] \begin{bmatrix} I_{r_1} \\ 0_{(n-r^s) \times r_1} \end{bmatrix} \begin{bmatrix} 0_{(r_1-1) \times r_j} \\ [\widehat{R}^1_{j,1}, \dots, \widehat{R}^1_{j,r_j}] \end{bmatrix} \\
&\qquad - [0_{(n-r^s) \times r_1}, I_{n-r^s}] \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} [I_{r_1}, 0_{r_1 \times (n-r^s)}] \begin{bmatrix} I_{r_1} \\ 0_{(n-r^s) \times r_1} \end{bmatrix} \begin{bmatrix} 0_{(r_1-1) \times r_j} \\ [\widehat{R}^1_{j,1}, \dots, \widehat{R}^1_{j,r_j}] \end{bmatrix} \\
&\quad \overset{(3.16)}{=} 0_{r_1 \times r_j} - [\mathcal{X}_{(n-r^s) \times (r_1-1)}, 0_{(n-r^s) \times 1}] \begin{bmatrix} 0_{(r_1-1) \times r_j} \\ [\widehat{R}^1_{j,1}, \dots, \widehat{R}^1_{j,r_j}] \end{bmatrix} = 0_{r_1 \times r_j},
\end{aligned} \tag{3.23}
$$

$$
\begin{aligned}
&\widehat{\mathcal{N}}_1 \begin{bmatrix} I_{r_1} \\ 0_{(n-r^s) \times r_1} \end{bmatrix} \widehat{S}_j \\
&\quad \overset{(3.11)}{=} [0_{(n-r^s) \times r_1}, I_{n-r^s}] [I_{r_1+n-r^s} - \widehat{\mathcal{B}}_1 (\widehat{\mathcal{C}}_1 \widehat{\mathcal{B}}_1)^{-1} \widehat{\mathcal{C}}_1] \begin{bmatrix} I_{r_1} \\ 0_{(n-r^s) \times r_1} \end{bmatrix} \begin{bmatrix} 0_{(r_1-1) \times r_j} \\ S^1 \end{bmatrix} \\
&\quad \overset{(3.23)}{=} 0_{r_1 \times r_j}
\end{aligned} \tag{3.24}
$$

and

$$
\begin{aligned}
&\widehat{\mathcal{N}}_1\left[\begin{bmatrix} I \\ 0 \end{bmatrix}\widehat{A}_{1,1}+\begin{bmatrix} 0 \\ I \end{bmatrix}\widehat{P}_1\right][I,0]\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}+\widehat{\mathcal{N}}_1\left[\begin{bmatrix} I \\ 0 \end{bmatrix}\widehat{S}_1+\begin{bmatrix} 0 \\ I \end{bmatrix}\widehat{Q}\right][0,I]\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad=\widehat{\mathcal{N}}_1\left(\begin{bmatrix} \widehat{A}_{1,1} \\ \widehat{P}_1 \end{bmatrix}[I,0]+\begin{bmatrix} \widehat{S}_1 \\ \widehat{Q} \end{bmatrix}[0,I]\right)\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad=\widehat{\mathcal{N}}_1\widehat{A}_1\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad=(\widehat{\mathcal{V}}_1\widehat{\mathcal{V}}_1^T)^{-1}\widehat{\mathcal{V}}_1^T[I_{r_1+n-r^s}-\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\widehat{\mathcal{C}}_1]\widehat{A}_1\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad=(\widehat{\mathcal{V}}_1\widehat{\mathcal{V}}_1^T)^{-1}\widehat{\mathcal{V}}_1^T[\widehat{A}_1\widehat{\mathcal{B}}_1-\widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\widehat{\mathcal{C}}_1\widehat{A}_1\widehat{\mathcal{B}}_1](\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad=(\widehat{\mathcal{V}}_1\widehat{\mathcal{V}}_1^T)^{-1}\widehat{\mathcal{V}}_1^T\Big[\big[(\widehat{\mathcal{B}}_1)_2^{(r_1+n-r^s)},\dots,(\widehat{\mathcal{B}}_1)_{r_1}^{(r_1+n-r^s)},*\big]\\
&\qquad-\ \widehat{\mathcal{B}}_1(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}[(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)_2^{(r_1)},\dots,(\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)_2^{(r_1)},*]\Big](\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad=\left(\widehat{\mathcal{V}}_1\widehat{\mathcal{V}}_1^T\right)^{-1}\widehat{\mathcal{V}}_1^T\left[\big[(\widehat{\mathcal{B}}_1)_2^{(r_1+n-r^s)},\dots,(\widehat{\mathcal{B}}_1)_{r_1}^{(r_1+n-r^s)},*\big]-\widehat{\mathcal{B}}_1\begin{bmatrix} 0 & * \\ I_{r_1-1} & * \end{bmatrix}\right](\widehat{\mathcal{C}}_1\widehat{\mathcal{B}}_1)^{-1}\\
&\quad\overset{(3.14)}{=}(\widehat{\mathcal{V}}_1\widehat{\mathcal{V}}_1^T)^{-1}\widehat{\mathcal{V}}_1^T[0,\dots,0,*]\begin{bmatrix} * & \cdots & * & 1 \\ \vdots & ⋰ & ⋰ & 0 \\ * & 1 & ⋰ & \vdots \\ 1 & 0 & \cdots & 0 \end{bmatrix}\\
&\quad=(\widehat{\mathcal{V}}_1\widehat{\mathcal{V}}_1^T)^{-1}\widehat{\mathcal{V}}_1^T[*,0,\dots,0].
\end{aligned}
\tag{3.25}
$$

Hence Eqs. (3.21)–(3.25) show that only the first r_1 columns of $\widehat{A}$ change when applying the transformation $\widehat{U}_1$. Furthermore the first r_1 columns of $\widehat{U}_1\widehat{A}\widehat{U}_1^{-1}$ are equal to the first r_1 columns of $\widetilde{A}$ and by (3.21)–(3.25) Eqs. (2.15) and (2.17) hold for $i=1$.

Moreover, an application of the transformation $\widehat{U}_i, i\in\{2,\dots,m\}$, has the similar effect as in (3.20)–(3.25) on the r_i columns from column number $\sum_{j=1}^{i-1}r_j+1$ to column number $\sum_{j=1}^{i}r_j$ of matrix $(\widehat{U}_{i-1}\ \cdots\ \widehat{U}_1\widehat{A}\widehat{U}_1^{-1}\ \cdots\ \widehat{U}_{i-1}^{-1})$, which, when finally all transformation matrices $\widehat{U}_i$ are applied, yields (2.2) and (2.15)–(2.18). This completes the proof. □

3.4. Proof of the zero dynamics

Now a proof for the characterization and stability of the zero dynamics of linear MIMO-systems with ordered vector relative degree is given.

Proof of Corollary 2.6. (i) Set

$$
\begin{aligned}
\mathcal{Z}=\{(\mathcal{V}\eta,-\Gamma^{-1}S\eta,0)\in\mathcal{C}^1([0,\infty)\to\mathbb{R}^n)\times\mathcal{C}^1([0,\infty)\to\mathbb{R}^m)\\
\times\,\mathcal{C}^1([0,\infty)\to\mathbb{R}^m)\ \Big|\ \dot\eta=Q\eta\}.
\end{aligned}
$$

"$\subseteq$": If $(x,u,y)\in\mathcal{ZD}(A,B,C)$ then $y\equiv 0$ on $[0,\infty)$ and so

$$
\xi=\left(y_1,y_1^{(1)},\dots,y_1^{(r_1-1)}\,\Big|\,y_2,\dots,y_2^{(r_2-1)}\Big|\cdots\Big|y_m,\dots,y_m^{(r_m-1)}\right)^T\equiv 0,
$$

which yields, in view of (2.1) and (2.2)

$$0_{m\times 1} = \underbrace{\begin{bmatrix} S^1 \\ \vdots \\ S^m \end{bmatrix}}_{=:S} \eta + \begin{bmatrix} c^1_{(n)} A^{r_1-1} B \\ \vdots \\ c^m_{(n)} A^{r_m-1} B \end{bmatrix} u, \quad \dot{\eta} = Q\eta,$$

thus (2.4) yields $u = -\Gamma^{-1} S\eta$. Since $x = U^{-1}(\xi^T, \eta^T)^T$ it follows from (3.3) and (3.19) that $(x, u, y) = (\mathcal{V}\eta, -\Gamma^{-1}S\eta, 0)$ for η being a solution of $\dot{\eta} = Q\eta$ and therefore, $(x, u, y) \in \mathcal{Z}$.
"$\supseteq$": Let $(\tilde{x}, \tilde{u}, \tilde{y}) = (\mathcal{V}\eta, -\Gamma^{-1}S\eta, 0) \in \mathcal{Z}$. By (2.7), $0 \equiv \tilde{y} = C\tilde{x} = C\mathcal{V}\eta$ thus

$$\tilde{\xi} = \left(\tilde{y}_1, \ldots, \tilde{y}_1^{(r_1-1)} \,\middle|\, \tilde{y}_2, \ldots, \tilde{y}_2^{(r_2-1)} \,\middle|\, \cdots \,\middle|\, \tilde{y}_m, \ldots, \tilde{y}_m^{(r_m-1)}\right)^T \equiv 0$$

and therefore $\left(\begin{pmatrix} 0 \\ \eta \end{pmatrix}, \tilde{u}, 0\right)$ solves (2.1), hence

$$(\tilde{x}, \tilde{u}, \tilde{y}) = \left(U^{-1}\begin{pmatrix} 0 \\ \eta \end{pmatrix}, \tilde{u}, 0\right) = (\mathcal{V}\eta, \tilde{u}, 0) \in \mathcal{Z}D(A, B, C).$$

(ii) From (i) it follows that

$$x = \mathcal{V}\eta \quad \text{and} \quad \eta = (\mathcal{V}^T\mathcal{V})^{-1}\mathcal{V}^T x,$$

where η is a solution of $\dot{\eta} = Q\eta$. Thus the zero dynamics of (A, B, C) are exponentially stable if, and only if, $\dot{\eta} = Q\eta$ is an exponentially stable system. $\square$

3.5. *Proof of the right-invertibility*

Proof of Corollary 2.7. "$\Rightarrow$": If $y = y_R$ then by (2.1) $\xi = \left(y_{R1}, \ldots, y_{R1}^{(r_1-1)} \,\middle|\, \cdots \,\middle|\, y_{Rm}, \ldots, y_{Rm}^{(r_m-1)}\right)^T$ and furthermore

$$y_{R_j}^{(r_j)} = y_j^{(r_j)} = \dot{\xi}_{\sum_{i=1}^j r_i} = R^j\xi + S^j\eta + c^j_{(m)}A^{r_j-1}Bu, \quad j \in \{1, \ldots, m\},$$

thus

$$\begin{bmatrix} y_{R1}^{(r_1)} \\ \vdots \\ y_{Rm}^{(r_m)} \end{bmatrix} = \begin{bmatrix} R^1 \\ \vdots \\ R^m \end{bmatrix} \xi + \begin{bmatrix} S^1 \\ \vdots \\ S^m \end{bmatrix} \eta + \underbrace{\begin{bmatrix} c^1_{(m)} A^{r_1-1} B \\ \vdots \\ c^m_{(m)} A^{r_m-1} B \end{bmatrix}}_{=\Gamma} u,$$

hence (2.19) and (2.20).

"$\Leftarrow$": Assume the (ii) holds. By (2.1) it follows that, for $j \in \{1, \ldots, m\}$

$$\dot{\xi}_{\sum_{i=1}^j r_i} = R^j\xi + S^j\eta + \underbrace{c^j_{(m)}A^{r_j-1}B\Gamma^{-1}}_{=e^j_{(m)}} \left(\begin{bmatrix} y_{R1}^{(r_1)} \\ \vdots \\ y_{Rm}^{(r_m)} \end{bmatrix} - \begin{bmatrix} R^1 \\ \vdots \\ R^m \end{bmatrix} \xi - \begin{bmatrix} S^1 \\ \vdots \\ S^m \end{bmatrix} \eta \right) = y_{R_j}^{(r_j)},$$

where for any $\eta^0 \in \mathbb{R}^{n-r^s}$, η is a solution of the initial value problem

$$\dot{\eta} = [P_1, \ldots, P_m]y_R + Q\eta, \quad \eta(0) = \eta^0.$$

This yields, in view of (2.1), $\xi = \left(y_{R1}, \ldots, y_{R1}^{(r_1-1)} \,\middle|\, \cdots \,\middle|\, y_{Rm}, \ldots, y_{Rm}^{(r_m-1)}\right)^T$ and thus $y = \Big(\xi_1, \xi_{r_1+1}, \ldots,$
$\xi_{\sum_{i=1}^{m-1} r_i+1}\Big)^T = y_R$ which completes the proof. $\square$

Acknowledgements

The author is indebted to Achim Ilchmann for his constant support and stimulation suggestions.

References

[1] C.I. Byrnes, A. Isidori, A frequency domain philosophy for nonlinear systems, with application to stabilization and to adaptive control, in: Proceedings of the 23rd IEEE Conference on Decision Control, vol. 1, 1984, pp. 1569–1573.
[2] C.I. Byrnes, A. Isidori, Global feedback stabilization of nonlinear systems, in: Proceedings of the 24th IEEE Conference on Decision Control, vol. 1, 1985, pp. 1031–1037.
[3] C.I. Byrnes, A. Isidori, Local stabilization of minimum-phase nonlinear systems,Systems Control Lett., 11 (1) (1988) 9–17.
[4] C.I. Byrnes, A. Isidori, New results and examples in nonlinear feedback, Systems Control Lett. 12 (5) (1989) 437–442.
[5] A. Ilchmann, M. Mueller, Time-varying linear systems: relative degree and normal form, IEEE Trans. Automat. Control 52 (5) (2007) 840–851.
[6] A. Ilchmann, E. Ryan, Universal λ-tracking for nonlinearly-perturbed systems in the presence of noise, Automatica 30 (2) (1994) 337–346.
[7] A. Ilchmann, E.P. Ryan, P. Townsend, Tracking with prescribed transient behavior for nonlinear systems of known relative degree, SIAM J. Control Optim. 46 (1) (2007) 210–230.
[8] A. Ilchmann, S. Townley, Simple adaptive stabilization of high-gain stabilizable systems,Systems Control Lett. 20 (3) (1993) 189–198.
[9] A. Isidori, Nonlinear control systems, Communications and Control Engineering Series, third ed., Springer-Verlag, Berlin, 1995.
[10] A. Isidori, Nonlinear control systems II, Communications and Control Engineering Series, Springer-Verlag, London, 1999.
[11] D. Liberzon, S.A. Morse, E.D. Sontag, Output–input stability and minimum-phase nonlinear systems, IEEE Trans. Automat. Control 47 (3) (2002) 422–436.
[12] S. Nicosia, A. Tornambè, High-gain observers in the state and parameter estimation of robots having elastic joints, Systems Control Lett. 13 (4) (1989) 331–337.
[13] R. Sepulchre, M. Janković, P.V. Kokotović, Constructive nonlinear control, Communications and Control Engineering Series, Springer-Verlag, Berlin, 1997.